Université de Montpellier
M2 Biostatistique - HAX918X

Année 2023-2024

Examen final - vendredi 27 octobre 2023
Durée 2h00 - Documents non autorisé

Exercice 1 (4 pts)

Trois individus arrivés indépendamment et au hasard à un arrêt de tram, se demandent quelle est la fréquence de passage des trams à cette période de la journée. Ils ont attendu 3, 4 et 6 minutes. Ils hésitent entre deux possibilités : les trams passent toutes les 8 minutes ou toutes les 12 minutes.

Nous supposons que les trois réalisations sont issues d'une loi uniforme sur $[0, \theta]$ avec $\theta \in \{8, 12\}$. En postulant une loi a priori uniforme sur les deux possibilités, $\mathbb{P}^\pi(\theta = 8) = \mathbb{P}^\pi(\theta = 12) = 1/2$, donner une résolution bayésienne de cette question.

Exercice 2 (8 pts)

Nous souhaitons estimer

$$I = \int_0^1 (1+x)^{-1/2} \exp(-x) \mathrm{d}x \,.$$

1 (2 pts) Proposer deux estimateurs de I par des méthodes de Monte Carlo différentes :

— la première stratégie sera basée sur une loi uniforme et l'estimateur correspondant sera noté $\hat{I}_N^1$ avec N le nombre de simulation ;

— la deuxième méthode sera basée sur une loi exponentielle et l'estimateur associé sera noté $\hat{I}_N^2$ avec N le nombre de simulation.

2 (2 pts) Comparer $\hat{I}_N^1$ et $\hat{I}_N^2$.

3 (2 pts) Donner un intervalle de confiance asymptotique de niveau 95% pour le paramètre I. Indication : utiliser le théorème central limite associé à $\hat{I}_N^1$.

4 (2 pts) En utilisant l'inégalité de Bienaymé-Tchebychev, déterminer le nombre de simulation nécessaire pour que l'erreur absolue relative de $\hat{I}_N^1$ soit inférieure à 1% avec une probabilité supérieure ou égale à 99%.

Exercice 3 (4 pts)

Nous considérons la variable aléatoire X admettant pour densité de probabilité

$$f_X(x) = \begin{cases} 4x & 0 \leq x \leq 1/2 \\ 4(1-x) & 1/2 < x \leq 1 \\ 0 & \text{sinon} \end{cases}$$

Proposer une méthode de simulation de réalisations suivant X.

Exercice 4 (4 pts)

Soit X une variable aléatoire telle que, $\forall x \in \mathbb{N}$,

$$\mathbb{P}(X = x|\theta) = \theta^x(1-\theta)$$

avec $\theta \in]0,1[$. Nous souhaitons estimer θ et nous nous plaçons dans le paradigme bayésien.

1) (2 pts) Donner la loi non informative de Jeffreys pour le paramètre θ.
Utiliser le fait que $\mathbb{E}(X|\theta) = \dfrac{\theta}{1-\theta}$.

2) (2 pts) Soit $(x_1, \ldots, x_n)$ la réalisation d'un n-échantillon de X. Donner la loi a posteriori de θ pour la loi a priori obtenue à la question précédente.

Correction Examen Final (1)
27 octobre 2023
HAX918X

Exercice 1

$X_1, X_2, X_3 | \theta \overset{iid}{\sim} \mathcal{U}_{[0,\theta]}$

$\Theta = \{8, 12\}$

$\mathbb{P}^{\pi}(\theta = 8) = \mathbb{P}^{\pi}(\theta = 12) = 1/2$

Calculons la loi posterior

$$\mathbb{P}^{\pi}(\theta = 8 | X_1, X_2, X_3) \propto \left(\frac{1}{2}\right) 8^{-3} \mathbb{1}_{[\max(X_1, X_2, X_3), +\infty[}(8)$$

$$\mathbb{P}^{\pi}(\theta = 12 | X_1, X_2, X_3) \propto \left(\frac{1}{2}\right) 12^{-3} \mathbb{1}_{[\max(X_1, X_2, X_3), +\infty[}(12)$$

$\max(X_1, X_2, X_3) = 6$

$$\Rightarrow \mathbb{P}^{\pi}(\theta = 8 | X_1, X_2, X_3) = \frac{8^{-3}}{8^{-3} + 12^{-3}}$$

$$\Longleftrightarrow \mathbb{P}^{\pi}(\theta = 12 \mid x_1, x_2, x_3) = \frac{12^{-3}}{8^{-3} + 12^{-3}}$$

Pour la fonction de perte symétrique nous choisissons comme estimation la valeur ayant la probabilité a posteriori la plus forte.

Comme $8^{-3} > 12^{-3}$ alors

$$\mathbb{P}^{\pi}(\theta = 8 \mid x_1, x_2, x_3) > \mathbb{P}^{\pi}(\theta = 12 \mid x_1, x_2, x_3)$$

$$\Rightarrow \hat{\theta}^{\pi} = 8 \quad \left(\mathbb{P}^{\pi}(\theta = 8 \mid x_1, x_2, x_3) \# 0{,}77 \right)$$

Exercice 2

1] $\hat{I}^1_N = \frac{1}{N} \sum_{i=1}^{N} (1 + X_i)^{-1/2} e^{-X_i}$

où $X_1, \ldots, X_N \overset{iid}{\sim} \mathcal{U}_{[0,1]}$

$\hat{I}^2_N = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{[0,1]}(X_i) \, (1 + X_i)^{-1/2}$

où $X_1, \ldots, X_N \overset{iid}{\sim} \mathcal{E}xp(1)$

2] $\mathbb{E}(\hat{I}^1_N) = \mathbb{E}(\hat{I}^2_N) = I$

Nous devons donc comparer $\mathbb{V}(\hat{I}^1_N)$ et $\mathbb{V}(\hat{I}^2_N$

3

$$\mathbb{V}(\hat{I}_N^1) = \frac{1}{N}\mathbb{V}_{\mathcal{U}_{[0,1]}}\left((1+X)^{-1/2}e^{-X}\right)$$

$$\mathbb{V}(\hat{I}_N^2) = \frac{1}{N}\mathbb{V}_{\mathcal{E}xp(1)}\left((1+X)^{-1/2}\mathbb{1}_{[0,1]}(X)\right)$$

Il s'agit donc de comparer

$$\mathbb{E}_{\mathcal{U}_{[0,1]}}\left((1+X)^{-1}e^{-2X}\right) = \int_0^1 (1+x)^{-1}e^{-2x}\,dx$$

$$\mathbb{E}_{\mathcal{E}xp(1)}\left((1+X)^{-1}\mathbb{1}_{[0,1]}(X)\right) = \int_0^1 (1+x)^{-1}e^{-x}\,dx$$

Comme $\forall x \in [0,1]$ $\quad e^{-2x} < e^{-x}$

alors $\mathbb{V}(\hat{I}_N^1) < \mathbb{V}(\hat{I}_N^2)$

3) $$\sqrt{N}\left(\hat{I}_N^1 - I\right) \underset{(N\ \text{grand})}{\approx} \mathcal{N}\Big(0, \underbrace{\mathbb{V}_{\mathcal{U}_{[0,1]}}\left((1+X)^{-1/2}e^{-X}\right)}_{\sigma}\Big)$$

$$\sigma = \mathbb{E}_{\mathcal{U}_{[0,1]}}\left((1+X)^{-1}e^{-2X}\right) - I^2$$

$$\le \mathbb{E}_{\mathcal{U}_{[0,1]}}\left(e^{-2X}\right) \quad \text{comme } X(\Omega) = [0,1]$$

$$= \int_0^1 e^{-2x}\,dx = \left[-\frac{e^{-2x}}{2}\right]_0^1 = \frac{1}{2} - \frac{e^{-2}}{2} < \frac{1}{2}$$

$$\hat{I}^1_N \pm \frac{\sqrt{v}}{\sqrt{N}} F^{-1}_{\mathcal{N}(0,1)}(0,975)$$

Intervalle de confiance asymptotique de niveau 95% !

Comme $v < \frac{1}{2}$, on peut utiliser sous-majorer la région asymptotique

$$\hat{I}^1_N \pm \frac{1}{\sqrt{2N}} F^{-1}_{\mathcal{N}(0,1)}(0,975)$$

4] D'après l'inégalité de Bienaymé-Tchebychev, nous avons

$$\mathbb{P}\left(\frac{|\hat{I}^1_N - I|}{|I|} \leq 0{,}01\right) \geq 1 - \frac{\mathbb{V}(\hat{I}^1_N)}{10^{-4} I^2}$$

Nous cherchons la plus petite valeur de N telle que $1 - \frac{\mathbb{V}(\hat{I}^1_N)}{10^{-4} I^2} \geq 0{,}99$

$$\frac{\mathbb{V}_{\mathcal{U}[0,1]}\left((1+x)^{-2} e^{-x}\right)}{N\, 10^{-4} I^2} \leq 0{,}01 \qquad N \geq \frac{10^6\, \mathbb{V}_{\mathcal{U}[0,1]}\left((1+x)^{-2} e^{-x}\right)}{I^2}$$

Comme $\forall_{[0,1]} \left|(1+x)^{-1} e^{-x}\right| < \frac{1}{2}$, (5)
nous cherchons la plus petite valeur de N telle que

$$N \geq \frac{10^6}{2 I^2}$$

Par ailleurs,

$$I = \int_0^1 (1+x)^{-1/2} e^{-x} dx$$

$$\frac{1}{e^x \sqrt{1+x}} \geq \frac{1}{\sqrt{2}\, e} \quad \forall x \in [0,1]$$

$$I \geq (2)^{-1/2} e^{-1}$$

$$I^2 \geq 2^{-1} e^{-2}$$

$$\frac{1}{I^2} \leq 2e^2$$

Nous cherchons donc la plus petite valeur de N telle que

$$N \geq 10^6 e^2$$

Il faut donc que $N \geq 7389057$

6

Exercice 3

Nous proposons d'utiliser un algorithme acceptation-rejet avec comme loi de proposition une loi uniforme sur [0,1]

$$\frac{f_X(x)}{g(x)} \leq 2 \quad \forall x \in [0,1]$$

Algorithme

Simuler $y \sim \mathcal{U}_{[0,1]}$ (*)

Calculer $\frac{f_X(y)}{2g(y)} = r$

Simuler $u \sim \mathcal{U}_{[0,1]}$, si $r \leq u$ alors $x = y$

sinon retourner à (*)

Exercice 4

1) $\log(f(x|\theta)) = x\log(\theta) + \log(1-\theta)$

$\forall x \in \mathbb{N}$

$$\frac{d \log f}{d\theta}(x|\theta) = \frac{x}{\theta} - \frac{1}{1-\theta}$$

$$\frac{d^2 \log f}{(d\theta)^2}(x|\theta) = -\frac{x}{\theta^2} - \frac{1}{(1-\theta)^2}$$

$$I_x(\theta) = \frac{1}{\theta(1-\theta)} + \frac{1}{(1-\theta)^2}$$

$$I_x(\theta) = \frac{1}{\theta(1-\theta)^2}$$

$$\Longrightarrow \pi^J(\theta) \propto \theta^{-1/2}(1-\theta)^{-1}\mathbb{1}_{]0,1[}(\theta)$$

loi impropre

2] $\pi^J(\theta|x_1,\dots,x_n) \propto \theta^{\sum_{i=1}^{n} x_i - 1/2}(1-\theta)^{n}\mathbb{1}_{]0,1[}(\theta)$

$$\Longrightarrow \theta|x_1,\dots,x_n \sim \text{Beta}\left(\sum_{i=1}^{n} x_i + 1/2, n\right)$$